

Exchange Rate Puzzles and Distorted Beliefs

Pierre-Olivier Gourinchas*

Princeton University, NBER and CEPR

Aaron Tornell

UCLA and NBER

6 June 2003

Abstract

We propose a new explanation for the foreign exchange forward-premium and delayed-overshooting puzzles. We show that both puzzles arise from a systematic distortion in investors's beliefs about the interest rate process. Accordingly, the forward premium is always a biased predictor of future depreciation; the bias can be so severe as to lead to negative coefficients in the 'Fama' regression. Delayed overshooting may or may not occur depending upon the persistence of interest rate innovations and the degree of misperception. We document empirically the extent of this distortion using survey data for G-7 countries against the U.S. and find that it is strong enough to account for these irregularities.

Keywords: Exchange Rates, Beliefs, Forward Premium Puzzle; Delayed Overshooting

JEL classification: E4, F31, G1.

1. Introduction

This paper proposes a new explanation for two important puzzles on foreign exchange behavior: the forward premium puzzle and the delayed overshooting puzzle. We demonstrate how both puzzles can arise from a specific distortion in beliefs about future interest rates. We document empirically the extent of this distortion for the G-7 countries and find that it is strong enough to explain these irregularities.

Over the past twenty five years, a substantial body of literature has documented the existence of large biases in the foreign exchange forward premium.¹ This 'Forward Premium Puzzle' (FPP) implies that nominal interest rate differentials between two countries bear little predictive power for the future rate of change in their nominal exchange rate. If anything, a *positive* forward premium is often associated with a subsequent *appreciation* of the exchange rate, not the *depreciation* that theory predicts. This empirical

*Corresponding author. E-mail: pog@princeton.edu. Phone: (609) 258-4026. Address: 311 Fisher Hall, Princeton University, Princeton NJ 08540. We thank Charles Engel and two anonymous referees as well as Daron Acemoglu, Olivier Blanchard, Ricardo Caballero, John Campbell, Rudi Dornbusch, Charles Engel, Jeff Frankel, Ben Friedman, Graciela Kaminski, John Leahy, Karen Lewis, Paul O'Connell, Julio Rotemberg, Nouriel Roubini and seminar participants at UCLA for useful discussions and comments. All errors are our own.

regularity implies predictable excess returns on foreign exchange markets.

A lesser known puzzle, the delayed overshooting puzzle, was documented by Eichenbaum and Evans (1995). These authors find that unanticipated contractionary shocks to U.S. monetary policy are followed by a *persistent* increase in U.S. interest rates and a *gradual appreciation* of the dollar, followed by a *gradual depreciation* several months later. This delayed overshooting pattern is consistent with predictable excess returns: for a while, U.S. interest rates are higher than foreign rates, *yet* the dollar appreciates, yielding positive excess returns. This dynamic pattern is also in contradiction with Dornbusch (1976)'s *overshooting* result, whereby the exchange rate should experience an immediate appreciation, and then depreciate gradually towards its new long run equilibrium value. While both puzzles are statements about predictable excess returns, they differ in subtle ways. The former is an *unconditional* statement about nominal interest rates and changes in exchange rates. The latter is a statement about their joint *conditional* response to a common unanticipated monetary innovation.²

In an accounting sense, there are two possible explanations for predictable excess returns: time-varying risk premia and/or expectational errors. Fully rational models with time-varying risk premia have had difficulties explaining these puzzles away.³ This paper allows instead for some departure from full-rationality. Our starting point is a setup where investors constantly try to determine the duration –transitory versus persistent– of interest rate shocks. In and by itself, this is not enough to account for the puzzles since rational agents cannot be systematically fooled. We thus introduce a particular distortion in the investors's beliefs about *future interest rates*. Specifically, we assume that investors misperceive the relative importance of transitory and persistent interest rate shocks, as measured by the variance of their innovations. Given these subjective beliefs, the exchange rate is determined by the standard no-arbitrage condition. In other words, subjective uncovered interest rate parity holds in our model.

We show that the model can exhibit both delayed overshooting and the forward premium bias if investors overestimate the importance of transitory shocks relative to persistent shocks, or equivalently, under-react to interest rate innovations. Moreover, for some parameter values, the forward premium bias can take its most extreme form, i.e. a negative correlation of the currency forward premium and the depreciation rate. Our model can also account for the volatility and the persistence puzzles observed in the data. Exchange rate changes are more variable than predictable excess returns, which in turn are more variable than the forward premium. Furthermore, exchange rate changes are less persistent than the forward premium.⁴

We document empirically the importance and direction of these distortions using two sources of interest rate expectations: monthly 3, 6 and 12 months ahead consensus forecast data on 3-month eurorates published by the *Financial Times Currency Forecaster*, and implicit forward rates constructed according to the rational expectation hypothesis of the term structure. In both cases, we compare the ‘true’ empirical process for the interest rate differential against the 3-month eurodollar, and the ‘subjective’ process that is consistent with the observed forecasts. We find substantial distortions in investors’s beliefs and strong evidence that investors overestimate the relative importance of transitory interest rate shocks. Strikingly, given our estimates of the subjective beliefs, our model is able to replicate the observed forward discount and delayed overshooting puzzles, as well as the volatility and the persistence puzzles.

To gain some intuition for our result, consider the following experiment. Suppose that U.S. interest rates increase relative to U.K. interest rates, then return gradually to their equilibrium value. If investors knew the exact nature of the interest rate shock, arbitrage would force the dollar to appreciate immediately relative to its long run value, up to the point where its expected future depreciation equals the forward premium. The dollar would then progressively revert to equilibrium as the forward premium vanishes. This is the *forward premium effect*. Suppose now that investors misperceive the U.S. interest rate shock as transitory. On impact, investors believe that the U.S. interest rate will revert to its equilibrium value fairly quickly. According to the subjective parity condition, the dollar need only appreciate moderately. In the following period, the U.S. interest rate turns out to be higher than investors first expected. This leads them to revise upward their beliefs about the persistence of the original interest rate shock and triggers a further appreciation of the dollar. We call this the *updating effect*. If the updating effect is strong enough to dominate the forward premium effect, there will be a gradual appreciation of the dollar. Eventually, there is not much more to learn and the forward premium effect must dominate. Thus, the exchange rate will revert to its equilibrium value. Along this path, there is delayed overshooting, positive excess returns in the domestic currency, and the forward premium is negatively correlated with expected appreciation.

It turns out that the same intuition carries over to an unconditional statement about the forward premium. Furthermore, while the model predicts that the forward premium bias always arises as soon as investors overestimate the importance of transitory shocks, the most extreme form of this puzzle (a negative correlation between the forward premium and the expected depreciation), and delayed overshooting, depend upon the parameters of the model and the relative strength of the two effects we just discussed. Intuitively then, hump-shaped exchange rate dynamics result from the interaction of the *misperception about the relative*

importance of interest rate shocks and the gradual mean reversion of interest rates.

We do not, in this paper, take a stand on the origins of the distorted beliefs. What we do is develop a *unified* model that allows us to link distorted beliefs about future interest rates to the two foreign exchange puzzles described above. The only deviation from full rationality in the model is the misperception about the relative variance of interest rate innovations. Given these beliefs, the exchange rate is determined optimally –and rationally. One important maintained assumption is that investors do not learn the true interest rate process over time. Why should they fail to adjust their beliefs? Clearly, with Bayesian learning, convergence to the true process would occur quite rapidly.⁵ However, this rapid learning is not compatible with the persistent distortions that we document. Something else must be at work. While we do not propose a full theory, we can offer a few remarks. First, it is possible that agents are rational, but econometricians condition upon a subset of the publicly available information. Such ‘Peso problem’ is consistent with our approach, although one would expect the biases to disappear as the sampling period extends.⁶ Another possibility is that while investors know the true model, they fear misspecification, and use ‘robust control’ to make forecasts and portfolio decisions. This leads them to act *as if* the approximating model has an extra fictitious sequence of transitory disturbances.⁷ Finally, we note that several recent papers have interpreted distorted beliefs as cognitive biases, and have used them to explain asset pricing anomalies.⁸ Unlike these papers, we carry out a direct empirical verification of the belief distortion that underlies our analysis.

Section 2 presents the model. Section 3 documents the empirical evidence on the objective and subjective processes for interest rates. Section 4 concludes. All proofs are relegated to Appendix A2..

2. Model

We start from the standard log-linearized foreign exchange arbitrage condition:

$$\mathcal{E}_t^s e_{t+1} - e_t = x_t - \zeta_t, \quad x_t := i_t - i_t^* \quad (1)$$

where e_t is the log of the domestic price of the foreign currency E_t , x_t is the forward premium, equal to the difference between the domestic one-period continuously compounded nominal interest rate i_t and its foreign equivalent i_t^* , and ζ_t is the domestic currency risk premium. $\mathcal{E}_t^s e_{t+1}$ represents the *subjective* expectation of next period’s exchange rate, which may differ from statistical or rational expectations, denoted by $\mathcal{E}_t e_{t+1}$. According to (1), the return on the short domestic bond, i_t , is equal to the return on a foreign bond of

the same maturity, i_t^* , adjusted for the market's expectations of depreciation $\mathcal{E}_t^s e_{t+1} - e_t$, as well as a risk premium component ζ_t . Appendix A2. presents an affine model of exchange rate determination consistent with (1).

We define *true* foreign exchange expected excess returns as the rationally expected excess return on the domestic currency:

$$\Lambda_t \equiv x_t - (\mathcal{E}_t e_{t+1} - e_t) = (\mathcal{E}_t^s e_{t+1} - \mathcal{E}_t e_{t+1}) + \zeta_t, \quad (2)$$

where the second expression substitutes for x_t using (1). According to (2), both expectation errors and risk premia generate true predictable excess returns. The difference between *subjective* expected excess returns, defined as $\Lambda_t^s \equiv x_t - (\mathcal{E}_t^s e_{t+1} - e_t) = \zeta_t$, and true expected excess returns is equal to the expectation error term $(\mathcal{E}_t^s e_{t+1} - \mathcal{E}_t e_{t+1})$.

Since subjective expectations and risk premia are not directly observed, the arbitrage relation (1) has no empirical power as it stands. If we assume that expectations are rational ($\mathcal{E}_t^s e_{t+1} = \mathcal{E}_t e_{t+1}$) and that there is no risk premium ($\zeta_t = 0$), then the (rationally) expected rate of depreciation $\mathcal{E}_t e_{t+1} - e_t$ is equal to the forward premium x_t , and there are no predictable excess returns, true or subjective ($\Lambda_t = \Lambda_t^s = 0$). This Uncovered Interest Parity (UIP) condition is the basis for all the empirical tests in the literature.

This paper focuses instead on expectational errors. Accordingly, we set the risk premium ζ_t to zero. It follows that the *subjective* expected rate of depreciation equals the forward premium:

$$\mathcal{E}_t^s e_{t+1} - e_t = x_t. \quad (3)$$

We will refer to (3) as ‘subjective UIP’. In that case, while there are no subjective predictable excess returns ($\Lambda_t^s = 0$), true predictable returns can differ from zero.

We take the forward premium x_t as the primitive object in our analysis. For simplicity, assume that x_t follows an autoregressive process with autocorrelation λ and an i.i.d. innovation ϵ_t normally distributed with mean zero and variance σ_ϵ^2 .⁹

$$x_t = \lambda x_{t-1} + \epsilon_t \quad (4)$$

While (4) characterizes the true process, we allow agents to have the following beliefs about the forward

premium process:

$$\begin{aligned} x_t &= z_t + v_t, \\ z_t &= \lambda z_{t-1} + \epsilon_t. \end{aligned} \tag{5}$$

The only difference between the true and subjective processes is the i.i.d. term v_t normally distributed with mean 0 and variance σ_v^{s2} . When $\sigma_v^{s2} = 0$, (5) collapses to (4). When $\sigma_v^{s2} > 0$, investors perceive interest rate changes as more transitory than they actually are. The difference between the true and subjective models is fully parameterized by the ratio of variances $\eta^s = \sigma_v^{s2}/\sigma_\epsilon^2$ that represents the subjective beliefs about the relative importance of transitory and persistent shocks. The true ratio satisfies $\eta = \sigma_v^2/\sigma_\epsilon^2 = 0$.

How should we interpret the shocks v_t and ϵ_t ? One possibility, consistent with sticky price models, is that v_t represents a relative velocity shock, and ϵ_t represents a permanent relative money supply shock. A permanent reduction in money supply leads to an increase in domestic interest rates. As prices adjust slowly over time, interest rates revert gradually toward their steady state value. This interpretation is consistent with the results of Eichenbaum and Evans (1995).¹⁰ Alternatively, the shocks v_t and ϵ_t can capture the uncertainty surrounding the conduct of monetary policy. Both the monetary policy target and the information set upon which Central Banks act are imperfectly known to the market.¹¹ Transitory shocks can occur when the Fed acts on inaccurate forecasts, or when the balance of power among the Open Market Committee members shifts temporarily. Both elements are not observed by market participants who have to infer the rationale behind the most recent policy decisions.

While investors act optimally conditional on their beliefs, we assume that they do not update these beliefs over time in order to learn the true model. In other words, it is never the case that eventually $\mathcal{E}_t^s x_{t+1} = \mathcal{E}_t x_{t+1}$. We offer three possible motivations for this assumption. First, this allows for a minimal departure from the full-rational model. In other words, since our agents are rational in every other way, we maintain a certain discipline. In particular, conditional on the perceived interest rate process (5), investors eliminate any apparent arbitrage opportunity, so that subjective UIP (3) holds. Second, distorted beliefs about σ_v^2 can be interpreted as a formalism to capture certain psychological biases as in behavioral finance.¹² Finally, suppose that agents know that (1)-(4) is indeed a good approximating model. However, they fear *misspecification* and in order to guard against it agents use ‘robust control’ to make forecasts and portfolio decisions. This leads them to act *as if* the approximating model has an extra fictitious sequence

of disturbances v_t . As a result, their forecasts might resemble those derived using (5) instead of (4). Here we take the misperception about σ_v^2 as a primitive. Our objective is to determine the extent to which this distortion can explain the forward premium puzzle and delayed overshooting.

2.1. Market Forecasts

Every period, agents observe the realization of the forward premium x_t and, given their beliefs (5), form forecasts of future forward premia ($\mathcal{E}_t^s x_{t+1}$) using Bayes law. The subjective interest parity condition (3) then determines the exchange rate.

This forecasting problem is a standard signal extraction problem. Its solution is given by the following lemma, which is a direct consequence of the properties of the Kalman Filter (see Hamilton (1994, chapter 13)).

Lemma 1 *Suppose that beliefs about z_1 are initially distributed as $\mathcal{N}(\mathcal{E}_0^s z_1, \sigma_1^2)$. Then:*

1. *market expectations of the forward premium evolve according to:*

$$\begin{aligned} \mathcal{E}_t^s x_{t+1} &= (1 - k_t) \lambda \mathcal{E}_{t-1}^s x_t + k_t \lambda x_t, \text{ where} \\ \sigma_{t+1}^2 &= (1 - k_t) (\lambda^2 \sigma_t^2 + \sigma_\epsilon^2), \quad 0 \leq k_t = \frac{\lambda^2 \sigma_t^2 + \sigma_\epsilon^2}{\lambda^2 \sigma_t^2 + \sigma_\epsilon^2 + \sigma_v^2} \leq 1 \end{aligned} \tag{6}$$

2. *in the limit as $t \rightarrow \infty$, the conditional variance σ_{t+1}^2 and the gain k_t converge to steady state values σ^2 and k respectively, that satisfy:*

$$k = \frac{1 + \Delta - \eta^s (1 + \lambda^2)}{1 + \Delta + \eta^s (1 + \lambda^2)}; \quad \sigma^2 = \frac{(1 - k)}{1 - (1 - k) \lambda^2} \sigma_\epsilon^2$$

where $\Delta^2 = [\eta^s (1 - \lambda^2) + 1]^2 + 4 \eta^s \lambda^2$, and the beliefs evolve according to:

$$\mathcal{E}_t^s x_{t+1} = (1 - k) \lambda \mathcal{E}_{t-1}^s x_t + k \lambda x_t \tag{7}$$

Forecasts of future forward premia $\mathcal{E}_t^s x_{t+1}$ are a weighted average of last period's forecast, $\mathcal{E}_{t-1}^s x_t$, and the current forward premium x_t . The gain k_t represents the weight given to current observations relative to past expectations while σ_{t+1}^2 represents the conditional variance of the market belief about the persistent component of the forward premium.

According to (6), when $\sigma_v^{s2} = 0$, full weight is given to the current forward premium ($k_t = 1$) and the forecast equation collapses to rational expectations: $\mathcal{E}_t^s x_{t+1} = \lambda x_t$. When $\sigma_v^{s2} > 0$, some weight is given to past forecasts ($k_t < 1$) and subjective expectations of future forward premia *under-react* to changes in x_t . It is well-known that (6) resembles updating rules under adaptive expectations. Muth (1960) has shown that if the underlying process can be represented as in (5), Bayesian updating takes the same form as adaptive expectations. The key difference is that under the former the weights depend on the relative variances of transitory and persistent components.

We assume that the updating process has been going on for long enough, so that the parameters $\{k_t, \sigma_{t+1}^2\}$ have converged to their steady state (k, σ^2) , and subjective beliefs satisfy (7).¹³ The steady state gain k depends only on the misperception coefficient (η^s) and the degree of persistence of the shocks (λ). It is easy to verify that the gain is increasing in λ and that it decreases from 1 to 0 as η^s varies from 0 to ∞ . Intuitively, with a higher λ , the current forward premium contains more information about future values of the persistent component z_t . Hence x_t gets more weight in the forecast. Given λ , there is a one-to-one mapping between the agent's misperception and the weight given to past beliefs. We can thus indifferently analyze the properties of the system in terms of (λ, η^s) or in terms of (λ, k) .

2.2. Equilibrium Exchange Rate

The equilibrium exchange rate is obtained by solving forward the subjective uncovered interest parity condition (3): $e_t = \sum_{j=0}^{T-1} \mathcal{E}_t^s x_{t+j} + \mathcal{E}_t^s e_{t+T}$. Substituting recursively using (6), we obtain the following proposition:

Proposition 2 *The equilibrium exchange rate satisfies:*

$$e_t = \bar{e}_t - x_t - \frac{1}{1-\lambda} \mathcal{E}_t^s x_{t+1} \quad (8)$$

where $\bar{e}_t = \lim_{T \rightarrow \infty} \mathcal{E}_t^s e_{t+T}$ is the long-run equilibrium value of the exchange rate.

The exchange rate depends upon the current forward premium x_t , its future expected value $\mathcal{E}_t^s x_{t+1}$ and the long-run equilibrium exchange rate \bar{e}_t .

Under rational expectations, agents do not misperceive a transitory component in the forward premium (i.e., $\mathcal{E}_t^s x_{t+1} = \lambda x_t$) and (8) simplifies to:

$$e_t^r = \bar{e}_t - \frac{x_t}{1 - \lambda} \quad (9)$$

where e_t^r denotes the exchange rate that would obtain without distorted beliefs. According to (9), an increase in domestic interest rates relative to foreign rates leads to an appreciation of the nominal exchange rate relative to its long run equilibrium. More persistent interest rate shocks require a larger initial appreciation of the currency in anticipation of larger future forward premia.

Subtracting (9) from (8), the equilibrium exchange rate can be expressed as the rational expectation exchange rate, plus a term that reflects the distortion in beliefs:

$$e_t = e_t^r + \frac{1}{1 - \lambda} (\mathcal{E}_t x_{t+1} - \mathcal{E}_t^s x_{t+1}) \quad (10)$$

With distorted beliefs, an increase in interest rate x_t implies that $\mathcal{E}_t x_{t+1} - \mathcal{E}_t^s x_{t+1} > 0$, since subjective forecasts underreact, so that the nominal exchange rate appreciates less than under rational expectations. Intuitively, since investors underestimate future domestic interest rates, the currency appreciates less initially.

2.3. Foreign Exchange Market Anomalies

This section characterizes the parameter set (λ, k) over which our simple model explains both Eichenbaum and Evans' delayed overshooting puzzle, as well as Fama's forward premium puzzle.

2.3.1. Predictable Excess Returns

In the absence of a risk premium, predictable excess returns on domestic currency originate exclusively from forecast errors. Iterating equation (2) forward, we obtain

$$\Lambda_t = (\bar{e}_t - \mathcal{E}_t \bar{e}_{t+1}) - \sum_{j=1}^{\infty} (\mathcal{E}_t^s x_{t+j} - \mathcal{E}_t \mathcal{E}_{t+1}^s x_{t+j}) \quad (11)$$

Predictable excess returns occur only if forecasts of the long run equilibrium exchange rate \bar{e}_t are incorrect, or if market expectations of interest rate differentials x_t differ from their statistical expectations, at least over some horizon.

While misperceptions about long run equilibrium values of the exchange rate may play some role in explaining predictable returns, we focus in this paper on predictable returns generated by distorted beliefs

about the interest rate. Consequently, we assume that $\mathcal{E}_t \bar{e}_{t+1} = \mathcal{E}_t^s \bar{e}_{t+1} = \bar{e}_t$.¹⁴ Substituting (6) into (11), we obtain the following result.

Proposition 3 . *There are true predictable excess returns as long as $\eta^s > 0$, and they satisfy:*

$$\Lambda_t = \left(1 + \frac{\lambda k}{1 - \lambda}\right) [\mathcal{E}_t x_{t+1} - \mathcal{E}_t^s x_{t+1}] \quad (12)$$

Positive predictable excess returns arise from forecast errors about future interest rates. The reason is simple: if future forward premia are under-estimated, the currency is artificially depreciated (see (10)) and will subsequently appreciate. When expectations are rational, $\Lambda_t = 0$, as expected.

2.3.2. Delayed Overshooting

The delayed overshooting path is characterized by Eichenbaum and Evans (1995) as the impulse response of the exchange rate to an unanticipated monetary shock. In the context of our model, this corresponds to the path that the exchange rate would follow if an innovation ϵ were to take place at time t , followed by no other shocks.¹⁵ Assume that the economy starts from steady state, with $\mathcal{E}_{t-1}^s x_t = \mathcal{E}_{t-1} x_t = 0$. According to (4), the true forward premium follows $x_{t+j} = \lambda^j \epsilon$. Under rational expectations, the exchange rate obeys (9) and satisfies:

$$e_{t+j}^r = \bar{e}_{t+j} - \frac{\lambda^j \epsilon}{1 - \lambda} \quad (13)$$

This path exhibits the standard Dornbusch (1976) overshooting result. Following an increase in domestic interest rates relative to foreign rates ($\epsilon > 0$), the exchange rate appreciates initially beyond its new equilibrium value, then depreciates gradually back at the same speed as the forward premium. We call this effect the **forward premium effect**.

By contrast, with distorted beliefs ($k < 1$), the response to a forward premium innovation ϵ at time t is:

$$e_{t+j} = \bar{e}_{t+j} - \frac{\lambda^j \epsilon}{1 - \lambda} [1 - \lambda(1 - k)^{j+1}] \quad (14)$$

The term in λ^j captures the forward premium effect as before. But there is another effect. Consider again the case of an increase in domestic interest rates relative to foreign interest rates. Initially, investors observe $x_t = \epsilon > 0$. They do not know whether this increase is persistent or transitory and form forecasts of x_{t+1} according to (7): $\mathcal{E}_t^s x_{t+1} = k\lambda\epsilon < \lambda\epsilon = \mathcal{E}_t x_{t+1}$. As we argued above, this leads to a smaller initial appreciation

of the exchange rate than under rational expectations. Consider now what happens at time $t + 1$. Investors observe $x_{t+1} = \lambda\epsilon$, which is higher than expected. They revise their beliefs according to (7) and estimate $\mathcal{E}_{t+1}^s x_{t+2} = \left(1 - (1 - k)^2\right) \lambda^2 \epsilon$. This revision may exceed the original estimate $k\lambda\epsilon$, leading investors to believe that the persistent component has increased. This **updating effect** leads to an appreciation of the exchange rate. The term $(1 - k)^{j+1}$ in equation (14) reflects this updating effect. Eventually, there is nothing to be learned and the forward premium effect must dominate. The humped exchange rate response results from the interaction between the forward premium and the updating effects.

Figure 1 shows the path of the exchange rate in response to an unanticipated increase in the forward premium at $t = 0$, when $k = 0.1$ and $\lambda = 0.98$. The exchange rate appreciates for about 15 periods before reverting back to its long run value. If one interprets each period as a month, this graph resembles the impulse response functions estimated by Clarida and Gali (1994), Eichenbaum and Evans (1995) and Grilli and Roubini (1994). Over what length of time does the exchange rate move in the ‘wrong’ direction? The next proposition characterizes delayed overshooting at horizon τ , i.e., the conditions under which $|e_{t+\tau} - \bar{e}_{t+\tau}| > |e_{t+\tau-1} - \bar{e}_{t+\tau-1}|$, as well as the ‘delayed overshooting region’ D_τ at horizon τ in terms of the parameters λ and k .

Proposition 4 . Delayed Overshooting:

1. *There is delayed overshooting at horizon τ if and only if $\eta > 0$ and:*

$$\tau < \frac{\ln \left(\frac{(1-\lambda)/\lambda}{1-\lambda(1-k)} \right)}{\ln(1-k)}; \quad (15)$$

2. *Delayed overshooting at horizon τ occurs if the parameters (λ, k) belong to the delayed overshooting region D_τ with lower boundary:*

$$\lambda(k, \tau) = \frac{1 + (1 - k)^\tau - \sqrt{\phi}}{2(1 - k)^{\tau+1}}$$

where $\phi = [1 + (1 - k)^\tau]^2 - 4(1 - k)^{\tau+1}$.

Panel A of figure 2 reports the lower boundary $\lambda(k, \tau)$ of D_τ as we vary τ from 1 to 10 periods. As we increase the peak date τ , the conditions on λ and k become more stringent: the frontier of D_τ shifts up. We see that higher λ makes delayed overshooting more likely, by slowing the forward premium effect. Changes in k have more complex effects. As k approaches 1, the updating process works more and more efficiently and beliefs converge to the true value very rapidly. In that case, the updating effect is dominated

by the forward premium effect and there is no delayed overshooting. Conversely, for sufficiently small k , updating occurs very slowly and forward premia observations convey little information about their persistent component. The subjective forecast $\mathcal{E}_t^s x_{t+1}$ does not increase much at the time of the shock, or following subsequent revisions. Delayed overshooting obtains for high persistence and low –but not too low– values of k .

Finally, true predictable excess returns on the domestic currency following an unanticipated increase in the forward premia satisfy:

$$\Lambda_{t+j} = \frac{\epsilon}{1-\lambda} \lambda^{j+1} (1-k)^{j+1} \geq 0 \quad (16)$$

Comparing (16) with (14) we observe that there are always positive true predictable excess returns on the domestic currency ($\Lambda_{t+j} \geq 0$) as soon as beliefs are distorted ($k < 1$), even if condition (15) is not satisfied.

2.3.3. The Forward Discount Puzzle

The previous discussion makes clear that there are conditional predictable excess returns. Are there also unconditional predictable returns, or equivalently, a forward premium puzzle? Under the null of Uncovered Interest Parity and rational expectations, the ‘Fama’ regression

$$e_{t+1} - e_t = \alpha + \beta x_t + u_{t+1} \quad (17)$$

yields a slope coefficient β equal to one. It is well known that these ‘Fama regressions’ fare quite badly. The estimated β , often called the ‘Fama coefficient’, is typically significantly smaller than one, and often negative. The asymptotic value of the OLS coefficient is equal to $\text{cov}(e_{t+1} - e_t, x_t) / \text{var}(x_t)$. In our model economy this asymptotic limit is given by the following proposition.

Proposition 5 . *The coefficient of the regression of realized depreciation rates $e_{t+1} - e_t$ on the forward premium x_t converges in plim to*

$$\beta_0 = 1 - \frac{\lambda(1-\lambda(1-k))(1-k)(1+\lambda)}{1-\lambda^2(1-k)} \leq 1 \quad (18)$$

One can verify directly that β_0 declines monotonically with λ . This is intuitive since a larger value of λ implies that any expectational error will lead to a more severe mispricing of the exchange rate. The dependence on k is more complex. A low β_0 requires a low k . However, when $k = 0$, which corresponds to an environment where agents believe all shocks are purely transitory ($\eta \rightarrow \infty$), β_0 is equal to $1 - \lambda > 0$. The

minimum of β_0 is attained for small but strictly positive values of k .

When $k < 1$ and $0 < \lambda \leq 1$, the asymptotic value of the Fama coefficient is strictly smaller than 1 and the forward premium is a biased predictor of the future rate of depreciation of the currency. But, equation (18) delivers a much stronger result. The second Panel B of figure 2 reports a contour plot of β_0 as function of λ and k . We see from the graph that it can be negative for large values of λ and small –but not too small– values of k .

It is easy to understand why β_0 must be smaller than 1. By definition, true expected depreciation is the difference between the forward premium and predictable excess returns: $\mathcal{E}_t e_{t+1} - e_t = x_t - \Lambda_t$. Using (12) and (6) to replace Λ_t by its equilibrium value, we obtain:

$$\mathcal{E}_t e_{t+1} - e_t = x_t - \lambda \left(1 + \frac{\lambda k}{1 - \lambda} \right) (1 - k) (x_t - \mathcal{E}_{t-1}^s x_t) \quad (19)$$

Let's fix $\mathcal{E}_{t-1}^s x_t$ for the time being, and consider an increase in the forward premium x_t at time t . Since $\lambda \left(1 + \frac{\lambda k}{1 - \lambda} \right) > 0$, expected depreciation responds by less than the forward premium. To understand what is going on, consider again an increase in domestic interest rates, relative to foreign rates. As we have seen, the currency initially appreciates less than under rational expectations. Hence, the *true* expected depreciation is *smaller* than that implied by the forward premium, and there are positive predictable excess returns. This is true whenever $0 < \lambda \leq 1$ and $\eta > 0$.¹⁶

Yet, we also found that β_0 can be negative, which is a much stronger result. How can this be? According to equation (19). If λ is large, the coefficient on the forward premium forecast error, $\lambda \left(1 + \frac{\lambda k}{1 - \lambda} \right) (1 - k)$, can *exceed* the coefficient on the forward premium (equal to 1), as predictable excess returns become more volatile. In that case, the initial mispricing of the exchange rate is so large that it requires the exchange rate to *appreciate* further in the future. As agents update their beliefs, they realize the change in interest rates is more persistent than initially anticipated. This upward revision in beliefs has a large effect on the exchange rate because agents expect high domestic interest rates to persist in the future. As a results we get an extreme scenario where a *high* domestic interest rate coexists with an *appreciating* currency. In other words, the forward premium and the true expected depreciation move in opposite directions. Thus, β_0 is negative.

The previous discussion abstracted from the term $\mathcal{E}_{t-1}^s x_t$. But past expected forward premia are correlated with the current forward premium. Indeed, one obtains (18) exactly when the correlation is properly taken

into account. This term dampens the movements in predictable excess returns and makes it more difficult for β_0 to turn negative (compare the term in x_t in (19) and the Fama coefficient).

To sum up, our analysis has strong cross-sectional implications. Countries should exhibit unconditional delayed overshooting and the forward discount puzzle in its most extreme form (i.e., a negative Fama coefficient), if (a) monetary shocks have high conditional persistence (high λ), and (b) the degree of misperception $(1 - k)$ is high, but not too high. Further, our analysis indicates that there are always positive predictable excess returns at short horizons ($\Lambda_t > 0$), even if there is no delayed overshooting.

2.3.4. The Volatility and Persistence Puzzles¹⁷

Backus, Gregory and Telmer (1993), Bekaert (1996) and Moore and Roche (2002) observe that exchange rate changes are many times more variable than predictable excess returns, which are in turn much more variable than the forward premium:

$$\text{var}(\Delta e_{t+1}) > \text{var}(\Lambda_t) > \text{var}(x_t) \quad (20)$$

This is the *volatility* puzzle. These authors also emphasize that the forward premium x_t is much more persistent than exchange rate changes Δe_{t+1} :

$$AC(\Delta e_t) > AC(x_t) \quad (21)$$

where $AC(x)$ denotes the autocorrelation of x . This is the *persistence* puzzle. A successful exchange rate model must also account for these volatility and persistence puzzles.

We compute the variances involved in (20) in the context of our misperception model, assuming that the long term exchange rate remains constant $\bar{e}_t = \bar{e} = 0$. Panel A of figure 3 reports the ratio of the unconditional volatility of predictable excess returns to the unconditional volatility of the forward premium ($\text{var}(\Lambda_t) / \text{var}(x_t)$), as a function of λ for various values of k . When $\lambda = 0$, there are no predictable excess returns, hence the ratio is zero. As λ increases, we see that the relative volatility increases, eventually exceeding one. Thus, for high values of λ , our model is able to account for the second inequality in the volatility puzzle (20).

Panel B reports the ratio of the unconditional volatility of the depreciation rate to the unconditional

volatility of the forward premium, $var(\Delta e_{t+1})/var(x_t)$. When $\lambda = 0$, the exchange rate satisfies $e_t = -x_t$, so that $var(\Delta e_{t+1})/var(x_t) = 2$. As λ increases, this ratio might increase or decrease, depending upon k . When $k = 0$, $e_t = -x_t$ so that $var(\Delta e_t)/var(x_t) = 2(1 - \lambda)$ decreases with λ . When $k = 1$, $e_t = -x_t/(1 - \lambda)$, and $var(\Delta e_t)/var(x_t) = 2/(1 - \lambda)$ can become very large for large λ . Importantly, for the values of λ and k that can explain the Fama puzzle and delayed overshooting, (high λ , small k), we can find large relative variances, so that our model can also account for the first inequality of the volatility puzzle (compare the scale of the top two panels of Figure 3). In sum, our model is able to reproduce the ranking in (20) and explain the volatility puzzle for high values of λ and low values of k .

To account for the persistence puzzle, we observe first that the true correlation of the forward premium is always equal to λ . Panel C of figure 3 reports the correlation of the rate of depreciation, $AC(\Delta e_t)$. This correlation increases with λ . For $\lambda = 0$, it is equal to -0.5 since $e_t = -x_t$. For $k = 0$ and $k = 1$ it is equal to $-(1 - \lambda)/2$. For intermediate values of k , the correlation can become positive, but remains always smaller than λ . Hence, the model can also account for the persistence puzzle (21).

3. What Do the Data Say, What Does the Market Think?

The simple model of the previous section accounts for all the puzzles as the consequence of a systematic and specific distortion in investor's beliefs about the interest rate process. This section finds strong evidence in support of this hypothesis. Our empirical strategy is as follows: we start by estimating a generalized version of the *true* empirical process according to (5) using 3-months eurorates for all G-7 eurorates against the 3-months eurodollar. In all cases, we find that $\eta = \sigma_v^2/\sigma_\epsilon^2 = 0$, while λ is close to 1. Then, using two different sources of interest rate forecasts, we estimate the *subjective* process (5) and show that η^s is significantly positive for most G-7 currencies against the U.S. dollar, and often larger than 1. We conclude that there is strong evidence of the type of subjective distortion that is necessary to explain the empirical puzzles. Finally, we show that the size of the bias induced by this belief distortion is sufficient to account for the empirical evidence on delayed overshooting, forward premium, persistence and volatility puzzles.

3.1. What Do the Data Say?

We start from a more general version of the empirical process for the forward premium of Section 2. Instead of equation (5), we consider:¹⁸

$$\begin{aligned} x_t &= z_t + v_t \\ \lambda(L) z_t &= \epsilon_t, \end{aligned} \tag{22}$$

where v_t is the transitory component and z_t is the persistent one, following an $AR(p)$ process with $\lambda(L) = 1 - \sum_{i=1}^p \lambda_i L^i$. The disturbances ϵ_t and v_t are independently and normally distributed with mean 0 and variance σ_v^2 and σ_ϵ^2 , respectively. Equation 22 provides a flexible parametric representation of the interest rate process. If $\sigma_v^2 = 0$, it is equivalent to an $AR(p)$; if $\sigma_\epsilon^2 = 0$, it reduces to a white noise process; finally, if the lag polynomial $\lambda(L)$ admits roots on the unit circle, x_t is integrated.

Equation (22) and the distributional assumptions on v_t and ϵ_t define a standard state-space process. The parameter vector $\theta_0 \equiv (\{\lambda_i\}_{i=1}^p, \sigma_\epsilon, \eta)$ can readily be estimated by Maximum Likelihood (see Hamilton (1994, chapter 13)).¹⁹

Our data set consists of monthly observations on the 3-months eurorates for Canada, France, Germany, Italy, Japan and the U.K., evaluated against the 3-months eurodollar. The sample period is 1986:8 to 1995:10 for a total of 111 observations.²⁰ Figure 4 reports the interest rate differential for all six countries against the U.S., over the sample period, together with the 3-months eurodollar. Table 1 presents Maximum Likelihood estimates of the relevant parameters for $p = 1$ to 4. A value of 0 for η indicates that the constraint $\eta \geq 0$ is binding, and the corresponding $AR(p)$ process, estimated directly, maximizes the likelihood.

First, we observe that the long run autocorrelation $\Sigma\lambda$ is very high, ranging from 0.95 for U.S.-Canada to 1.01 for U.S.-Japan. Second, there is no evidence of a transitory component in the forward premium. More specifically, in all cases except U.S.-Canada, the constraint $\eta \geq 0$ binds, implying $\sigma_v^2 = 0$.²¹ These results imply that the forward premium is best characterized as following a very persistent AR process. In the AR(1) case, this corresponds to the process (4).

Since we find very persistent autocorrelations, Table 1 reports a Phillips-Perron test of non-stationarity for the forward premium (See Hamilton (1994, chapter 17)). The statistic Z_t tests the null ($\alpha = 0, \beta = 1$) of the regression $x_t = \alpha + \beta x_{t-1} + \varepsilon_t$.²² The table reports the associated p -value. The results indicate that

we can reject the null for all pairs at conventional levels of significance, except U.S.-Italy and U.S.-Germany. Although the forward premium is very persistent, it does not appear to be integrated.

3.2. What Does the Market Think?

Part I: Consensus Forecasts

Given an initial estimate for the persistent component $\mathcal{E}_0 z_1$, we can use equation (7) to construct forecasts of the forward premium at horizon τ , for a given set of subjective beliefs summarized by the parameter vector $\theta^s = (\{\lambda_i^s\}_{i=1}^p, \sigma_\epsilon^s, \eta^s)$.²³ Consider for instance the AR(1) case: $z_t = \lambda^s z_{t-1} + \epsilon_t$. The τ -period ahead forecast of the forward premium satisfies: $\mathcal{E}_t^s x_{t+\tau} = \mathcal{E}_t^s z_{t+\tau} = (\lambda^s)^{\tau-1} \mathcal{E}_t^s z_{t+1}$. Using equation (7) and iterating back to the beginning of the sample, we can write:

$$\mathcal{E}_t^s x_{t+\tau} = (\lambda^s)^{\tau-1} \left(k (\lambda^s) \sum_{j=0}^{t-1} (\lambda^s)^j (1-k)^j x_{t-j} + (1-k)^t \lambda^t \mathcal{E}_0 z_1 \right) \equiv \mathcal{E}_t (x_{t+\tau} | \theta^s) \quad (23)$$

The forecasts $\mathcal{E}_t^s x_{t+\tau}$ are a function of θ^s , the initial belief $\mathcal{E}_0 z_1$, and the history of forward premia up to time t , $\{x_s\}_{s=1}^t$. If we assume that forecasts are formed according to (23), we can estimate the parameter vector θ^s that is consistent with these forecasts and the forward premium process.²⁴ Specifically, suppose that we observe forecast data at time t and horizon τ , x_t^τ , with some classical measurement error u_t :

$$x_t^\tau = \mathcal{E}_t (x_{t+\tau} | \theta^s) + u_t \quad (24)$$

Under the assumption that u_t is normally distributed with mean 0, variance σ_u^2 , and uncorrelated with the true forecast, we can estimate θ^s by Maximum Likelihood.

We construct the empirical counterpart of x_t^τ using Eurorates consensus forecasts at 3, 6 and 12 months from the *Financial Times Currency Forecaster*. The data is available monthly from August 1986 to October 1995. Contributors include multinational companies as well as forecasting services from major investment banks, i.e., the most active players on the fixed income and foreign exchange markets.²⁵ The monthly publication collects interest rates and their forecasts and reports a ‘market average’ weighting individual forecasts according to their relative importance. We construct forecasts of the forward premium x_t by simple difference of the consensus forecasts of the underlying interest rates i_t^τ : $x_{t+\tau}^\tau = i_{t+\tau}^{us,\tau} - i_{t+\tau}^{*,\tau}$. Figure 4 reports the 3, 6 and 12 months forecasts for all country pairs against the U.S. dollar.

A few remarks are in order when using survey data. First, while there is substantial heterogeneity in forecasts, ‘market expectations’ may possess better statistical properties than individual forecasts, as the idiosyncratic noise ‘washes out’ in the aggregation process. Second, a recent theoretical literature has emphasized that there may be systematic biases in individual forecasts: forecasters who care about their reputation, may have incentives to use forecasts in order to manipulate their clients’s belief regarding their ability. However, the empirical importance and direction of such reputational biases remains an open question.²⁶ While we recognize the importance of these considerations, we do not find any compelling reason to believe that consensus forecast are systematically biased and we maintain the assumption that u_t is uncorrelated with market forecasts.

Table 2 presents the results when we pool all forecast horizons.²⁷ The estimates of the long run autocorrelation $\Sigma\lambda^s$ are very similar to the original estimates ranging from 0.95 to 1.00. More importantly, we find significant estimates of the relative variance η^s for most pairs, except U.S.-Germany (AR(1)) and U.S.-Canada (AR(3)). Since we have seen that $\eta = 0$ in the data generating process, a positive estimate of η^s implies that investors systematically misperceive interest rate shocks to be more transitory than what they actually are. Note also that in many cases $\eta^s > 1$. This indicates that investors believe –erroneously– that transitory shocks represent the largest fraction of interest rate innovations.

Part II: Using the Term Structure

As an alternative to consensus forecasts, we construct interest rate forecasts from the forward rates implicit in the term structure. Denote i_t^n the continuously compounded annualized yield on a n –period euro deposit.²⁸ According to the log-linearized version of the term structure, the n -months ahead k -periods forward rate at time t , denoted $f_{n,t}^k$, is equal to $f_{n,t}^k = \left(1 + \frac{n}{k}\right) i_t^{n+k} - \frac{n}{k} i_t^n$.²⁹ Under the rational expectations hypothesis of the term structure, this forward rate is an unbiased predictor of the k -months interest rate at time $t + n$: $f_{n,t}^k = \mathcal{E}_t i_{t+n}^k$. We use $f_{n,t}^k$ in place of the consensus forecast $\mathcal{E}_t^s i_{t+n}^k$ to estimate θ^s .

We use monthly data on the 3, 6 and 9 months eurorates for all G-7 countries to construct the 3 and 6 months-ahead forward premium on the 3-month eurorate. Table 3 presents the Maximum Likelihood estimates of θ^s analogous to table 2, for $p = 1$ to 4.

While the results are slightly weaker than using consensus forecasts, we still find strong evidence of misperception for $p = 2$ to $p = 4$. For all country-pairs, except U.S.-Canada ($p = 2$ and 4), and U.S.-Italy ($p = 3$), we estimate $\eta^s > 0$. In many instances, we also find that $\eta^s > 1$.³⁰

3.3. *Is it Big Enough?*

The previous subsection established that $\eta^s > 0$ for most country pairs against the U.S. As Section 2 makes clear, $\eta^s > 0$ is a necessary, but not a sufficient, condition for the exchange rate puzzles that we seek to explain. To illustrate how much of these puzzles our simple model is able to account for, we report in Table 4 estimates of the gain of the filter k , the delayed overshooting horizon τ , the Fama coefficient β , and the volatility and persistence ratios for the AR(2) estimates of Table 2. The first line for each country reports the empirical estimates from the data.³¹

We find that the model is qualitatively able to reproduce *all* stylized facts. The gain k ranges between 0.27 (U.S.-Italy) and 0.68 (U.S.-Canada), implying substantial amounts of underreaction. Based on the degree of misperception, we would predict a significant delayed overshooting ranging from 2.25 months (U.S.-Canada) to 13.45 months (U.S.-Italy), lower than the Eichenbaum and Evans (1995) estimates, ranging from 24 months to 38 months. Our model predicts significantly negative estimates for all exchange rates except U.S.-Canada. This is in line with empirical estimates of the Fama regression except the for the dollar-Lira.

We find also large excess volatility of the rate of depreciation compared to the forward premium. The empirical estimates range from 30.02 (U.S.-Canada) to 109.65 (U.S.-Japan), while the implied estimates range from 20 (U.S.-Canada) to 80.4 (U.S.-Italy). The model also implies that predictable excess returns Λ_t are more volatile than the forward premium x_t , although the ratio is often larger than the empirical estimates.

Finally, the model-implied autocorrelation of the rate of depreciation $AC(\Delta e_t)$, between 0.25 and 0.72, are always lower than the autocorrelation of the forward premium (λ , from table 2), but larger than the empirical counterpart, from -0.06 to 0.08.

Overall, we find these results quite striking, especially if one keeps in mind that we only use information about the subjective interest rate process to explain foreign exchange market puzzles. We conclude that, despite its simplicity, our model captures salient features of exchange rate behavior.

4. Conclusion

In this paper we establish a link between exchange rate and interest rate anomalies. We develop a nominal exchange rate determination model in which investors constantly try to determine whether interest rate shocks are transitory or persistent. We demonstrate that if investors misperceive shocks to be more

transitory than what they actually are, the equilibrium exchange rate in the model can account for four anomalies present in the data: the forward premium puzzle in its most extreme form – a negative Fama coefficient, delayed overshooting, as well as the variance and persistence puzzles. We then show that the data strongly supports the existence of this misperception. Furthermore, we show that for the degree of misperception that we estimate, we can account for a large portion of the puzzles.

The dynamics of the exchange rate are determined by two effects which act in opposite directions: a forward premium effect and a updating effect. The first effect is standard. As in other models, it derives from the mean reverting nature of interest rates differentials. The updating effect is novel, and derives from the fact that agents constantly try to determine whether interest rate shocks are transitory or persistent. The predictable excess returns and hump-shaped dynamics, implied by the currency market anomalies, can result if and only if agents misperceive interest rate shocks to be more transitory than what they actually are.

This interpretation, which is new to our knowledge, has important implications. It provides a clear analytical characterization of the factors influencing exchange rate responses to monetary shocks. It also serves the useful purpose of uncovering a deeper rationale for a variety of asset market pathologies. In particular, the misperceptions we identify raise the interesting prospect of an integrated understanding of currency and bond markets.

Of course, this paper also raises some intriguing questions: why do agents fail to revise their erroneous beliefs about second moments? Can these misperceptions be arbitrated away or taken advantage of by savvier investors? Ultimately, of course, we will need to reconcile observed behavior with models of optimal behavior. While we may not be there yet, this paper indicates a promising avenue of research.

References

- Abel, Andrew**, “An Exploration of the Effects of Pessimism and Doubt on Asset Returns,” *Journal of Economic Dynamics and Control*, 2002, 26 (7-8), 1075–1092.
- Backus, David K., Allan Gregory, and Christopher I. Telmer**, “Accounting for Forward Rates in Markets for Foreign Currency,” *Journal of Finance*, 1993, 58, 1887–1908.
- Backus, David, Silverio Foresi, and Chris Telmer**, “Affine Models of Currency Pricing: Accounting for the Forward Premium Anomaly,” July 1998. mimeo New York University.
- Barberis, Nicholas, Andrei Shleifer, and Robert Vishny**, “A Model of Investor Sentiment,” *Journal of Financial Economics*, September 1998, 49 (3), 307–43.
- Bekaert, Geert**, “The Time-Variation of Risk and Return in Foreign Exchange Markets: A General Equilibrium Perspective,” *Review of Financial Studies*, 1996, 9, 427–440.
- Cecchetti, Stephen G., Pok-Sang Lam, and Nelson C. Mark**, “Asset Pricing with Distorted Beliefs, Are Equity Returns Too Good to Be True?,” *American Economic Review*, 2000, 90 (4), 787–805.
- Clarida, Richard and Jordi Gali**, “Sources of Real Exchange Rate Fluctuations: How Important Are Nominal Shocks?,” 1994. NBER Working Paper 4658.
- Dornbusch, Rudiger**, “Expectations and Exchange Rate Dynamics,” *Journal of Political Economy*, 1976, 84, 1161–1176.
- Ehrbeck, Tilman and Robert Waldmann**, “Why Are Professional Forecasters Biased? Agency Versus Behavioral Explanations,” *Quarterly Journal of Economics*, Feb 1996, 111 (1), 21–40.
- Eichenbaum, Martin and Charles Evans**, “Some Empirical Evidence on the Effects of Monetary Policy Shocks on Exchange Rates,” *Quarterly Journal of Economics*, Nov 1995, 110, 975–1009.
- Frankel, Jeffrey and Andy Rose**, “Empirical Research on Nominal Exchange Rates,” in Gene M. Grossman and Kenneth Rogoff, eds., *Handbook of International Economics*, Amsterdam: North Holland, 1995, pp. 1689–1729. NBER Working Paper 4865.
- **and Kenneth Froot**, “Forward Discount Bias: Is it an Exchange Risk Premium?,” *Quarterly Journal of Economics*, 1989, 104, 139–161.

- Froot, Kenneth**, “New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates,” *Journal of Finance*, 1989, *44* (2), 283–304.
- Gourinchas, Pierre-Olivier and Aaron Tornell**, “Exchange Rate Dynamics, Learning and Misperception,” December 2002. NBER Working Paper 9391.
- Grilli, Vittorio and Nouriel Roubini**, “Liquidity and Exchange Rates: Puzzling Evidence From the G-7 Countries,” 1994. mimeo Yale University.
- Hamilton, James**, *Time Series Analysis*, Princeton University Press, 1994.
- Hansen, Lars, Thomas Sargent, and Thomas Tallarini**, “Robust Permanent Income and Pricing,” *Review of Economic Studies*, 1999, *66* (4), 873–907.
- Hodrick, Robert**, “The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets,” in “Fundamentals of Pure and Applied Economics” Harwood Academic Publishers 1988.
- Hong, Harrison and Jeremy Stein**, “A Unified Theory of Underreaction, Momentum Trading and Overreaction in Asset Markets,” *Journal of Finance*, December 1999, *54* (6), 2143–84.
- Kaminski, Graciela**, “Is There a Peso Problem? Evidence form the Dollar/Pound Exchange Rate, 1976–1987,” *American Economic Review*, 1993, *83*, 450–472.
- Lewis, Karen**, “Changing Beliefs and Systematic Rational Forecast Errors with Evidence from Foreign Exchange,” *American Economic Review*, 1989, *79*, 621–636.
- , “Puzzles in International Financial Markets,” in Gene M. Grossman and Kenneth Rogoff, eds., *Handbook of International Economics*, North Holland Elsevier Amsterdam 1995, pp. 1913–1972.
- Moore, Michael J. and Maurice J. Roche**, “Less of a Puzzle: A New Look at the Forward Forex Market,” *Journal of International Economics*, 2002, *58*, 387–411.
- Muth, John**, “Optimal Properties of Exponentially Weighted Forecasts,” *Journal of the American Statistical Association*, 1960, *55*, 299–306.
- Shiller, Robert, John Campbell, and K.L. Schoenholtz**, “Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates,” *Brookings Papers on Economic Activity*, 1983, pp. 223–242.
- Tornell, Aaron**, “Exchange Rate Anomalies Under Model Misspecification: A Mixed Optimal/Robust Approach,” 2002. UCLA mimeo.

A Appendices

A1. Data sources

Variable	Definition	Database
Euro rates	average of the 3 months eurocurrency deposit bid and ask rate	DRI/Reuters
Consensus Forecasts	3,6 and 12 months consensus forecasts on 3 months Euro rates	Financial Times Currency Forecaster

A2. An Affine Model of Exchange Rate Determination

This appendix presents a simple affine model of exchange rate determination under complete markets, similar to Backus, Foresi and Telmer (1998). The economy consists of two countries, each with its own currency. In each country, nominal assets are traded. The absence of pure arbitrage opportunities implies the existence of [a] positive nominal pricing kernels for claims in domestic and foreign currency, $M_{t,t+1}$ and $M_{t,t+1}^*$ respectively, such that:

$$1 = \mathcal{E}_t^s \{M_{t,t+1} R_{t+1}\}, \quad 1 = \mathcal{E}_t^s \{M_{t,t+1}^* R_{t+1}^*\}, \quad (\text{A.1})$$

where $R_{t+1}(R_{t+1}^*)$ is the gross rate of return -in domestic(foreign) currency- on any traded asset between t and $t+1$. Using these pricing kernels, the one-period continuously compounded nominal risk-free rates satisfy:

$$i_t = -\log \mathcal{E}_t^s \{M_{t,t+1}\}; \quad i_t^* = -\log \mathcal{E}_t^s \{M_{t,t+1}^*\} \quad (\text{A.2})$$

Let E_t denote the domestic price of foreign currency (so that an increase in E_t corresponds to a depreciation of the domestic currency). Under complete markets, and with identical preferences across countries, the nominal pricing kernel is unique and such that:

$$M_{t,t+1} \frac{E_{t+1}}{E_t} = M_{t,t+1}^* \quad (\text{A.3})$$

This equation pins down the rate of depreciation, not the level of the exchange rate. Taking logs, letting $e_t = \log(E_t)$, and substituting for the domestic and foreign nominal interest rates gives a general form of the foreign exchange arbitrage condition:

$$\mathcal{E}_t^s e_{t+1} - e_t = x_t + \zeta_t, \quad (\text{A.4})$$

where $x_t = i_t - i_t^*$

and $\zeta_t = (\mathcal{E}_t^s \log M_{t,t+1}^* - \log \mathcal{E}_t^s \{M_{t,t+1}^*\}) - (\mathcal{E}_t^s \log M_{t,t+1} - \log \mathcal{E}_t^s \{M_{t,t+1}\})$

We assume that the pricing kernels follow log-normal processes:

$$\ln M_{t,t+1} = -\ln R - \frac{\bar{\varphi}^2 \sigma^2}{2} - \delta \bar{z}_t - \frac{\varphi^2 \sigma^2}{2} - z_t - \bar{\varphi} \bar{\epsilon}_{t+1} - \varphi \epsilon_{t+1} \quad (\text{A.5})$$

$$\ln M_{t,t+1}^* = -\ln R - \frac{\bar{\varphi}^{*2} \sigma^2}{2} - \delta^* \bar{z}_t - \frac{\varphi^{*2} \sigma^2}{2} - z_t^* - \bar{\varphi}^* \bar{\epsilon}_{t+1} - \varphi^* \epsilon_{t+1}^*, \quad (\text{A.6})$$

where the elements of $\epsilon_{t+1} = (\bar{\epsilon}_{t+1}, \epsilon_{t+1}, \epsilon_{t+1}^*)'$ are independent and normally distributed with mean 0 and variance σ_ϵ^2 .³² The positive parameters $\delta, \delta^*, \varphi, \varphi^*, \bar{\varphi}$ and $\bar{\varphi}^*$ represent the *loadings* on the various shocks; R is a strictly positive scalar; \bar{z}_t and $\bar{\epsilon}_{t+1}$ represent the predictable and unpredictable components of a shock common to both countries, while (z_t, ϵ_{t+1}) and $(z_t^*, \epsilon_{t+1}^*)$ represent the predictable and unpredictable components of country-specific shocks.

We assume that the state $\mathbf{z}_t = (\bar{z}_t, z_t, z_t^*)'$ obeys the following AR process:

$$\lambda(L) \mathbf{z}_{t+1} = \epsilon_{t+1} \quad (\text{A.7})$$

where $\lambda(L)$ is the polynomial in the lag operator of order $p : 1 - \sum_{i=1}^p \lambda_i L^i$. Assumption (A.7) imposes that each component of \mathbf{z}_{t+1} has the same autocorrelogram. This can easily be relaxed. From (A.5)-(A.6), the interest rate and depreciation rates satisfy:

$$x_t = (\delta - \delta^*) \tilde{z}_t + z_t - z_t^* \equiv \tilde{z}_t \quad (\text{A.8})$$

$$\mathcal{E}_t^s \{e_{t+1}\} - e_t = x_t + \zeta \quad (\text{A.9})$$

The forward premium is a function of \tilde{z}_t only. \tilde{z}_t follows an AR(p) with innovation $\tilde{\epsilon}_t = (\delta - \delta^*) \bar{\epsilon}_t + \epsilon_t - \epsilon_t^*$ and variance $\tilde{\sigma}_\epsilon^2 = [(\delta - \delta^*)^2 + 2] \sigma_\epsilon^2$. The risk premium ζ is constant and equal to $(\bar{\varphi}^2 - \bar{\varphi}^{*2} + \varphi^2 - \varphi^{*2}) \sigma_\epsilon^2 / 2$. It depends upon the variance of the innovations and the loading factors. Since this risk premium is not time-varying, it is irrelevant for our analysis. We set it to 0 by assuming $\bar{\varphi} = \bar{\varphi}^*$ and $\varphi = \varphi^*$, implying that *subjective uncovered interest parity (3) holds exactly*.

The previous assumptions imply that x_t follows an AR(p) process:

$$\lambda(L) x_t = \lambda(L) \tilde{z}_t = \tilde{\epsilon}_t \quad (\text{A.10})$$

A2.1. Distorted Beliefs

Assume that investors assume that the pricing kernels satisfy:

$$\ln M_{t,t+1}^s = \ln M_{t,t+1} - v_t, \quad \ln M_{t,t+1}^{*,s} = \ln M_{t,t+1}^* - v_t^* \quad (\text{A.11})$$

where v_t and v_t^* are independent normally distributed shocks with mean 0 and variance $\sigma_v^2/2$. Interest and exchange rates are still determined by (A.2), and the exchange rate still satisfies (A.3), with M^s and $M^{*,s}$ in place of M and M^* .

According to (A.2), the forward premium satisfies:

$$x_t = \tilde{z}_t + \tilde{v}_t \quad (\text{A.12})$$

where $\tilde{v}_t = v_t - v_t^*$. Equation (A.12) generalizes equation (5) in Section 2 since \tilde{z}_t follows an AR(p).

A2.2. Updating

Given their beliefs, agents form forecasts of future forward premia using Bayes law. The system (x_t, \tilde{z}_t) defines a state-space system. We rewrite it in canonical form as follows:

$$\begin{aligned} x_t &= \mathbf{H}' \tilde{\mathbf{z}}_t + \tilde{v}_t \\ \tilde{\mathbf{z}}_t &= \mathbf{F} \tilde{\mathbf{z}}_{t-1} + \tilde{\epsilon}_t \end{aligned}$$

where $\tilde{\mathbf{z}}_t = (\tilde{z}_t, \tilde{z}_{t-1}, \dots, \tilde{z}_{t-p+1})'$, $\tilde{\epsilon}_t = (\tilde{\epsilon}_t, 0, \dots, 0)$, $\mathbf{H}' = (1, 0, \dots, 0)$ and \mathbf{F} is defined so that (A.10) holds (when $p = 1$, $\mathbf{F} = \lambda$). The following lemma is a direct application of chapter 13 in Hamilton (1994) and characterizes the evolution of these subjective beliefs.

Lemma 6 Assume that beliefs about $\tilde{\mathbf{z}}_1$ are initially distributed as $\mathcal{N}(\mathcal{E}_1^s \tilde{\mathbf{z}}_{1|0}, \mathbf{P}_1)$. Then:

1. Beliefs evolve according to:

$$\mathcal{E}_t^s \tilde{\mathbf{z}}_{t+1} = \mathbf{F} \mathcal{E}_{t-1}^s \tilde{\mathbf{z}}_t + \mathbf{F} \mathbf{P}_t \mathbf{H} (\mathbf{H}' \mathbf{P}_t \mathbf{H} + \sigma_v^{s2} \mathbf{I})^{-1} (x_t - \mathbf{H}' \mathcal{E}_{t-1}^s \tilde{\mathbf{z}}_t)$$

2. The conditional variance \mathbf{P}_t evolves according to:

$$\mathbf{P}_{t+1} = \mathbf{F} \left[\mathbf{P}_t - \mathbf{P}_t \mathbf{H} (\mathbf{H}' \mathbf{P}_t \mathbf{H} + \sigma_v^{s2} \mathbf{I})^{-1} \mathbf{H}' \mathbf{P}_t \right] \mathbf{F}' + \tilde{\sigma}_\epsilon^2 \mathbf{I}$$

in particular, it does not depend upon the actual realizations of the forward premium x_t .

3. In the limit as $t \rightarrow \infty$, the conditional variance converges to a steady state value \mathbf{P} , solution of:

$$\mathbf{P} = \mathbf{F} \left[\mathbf{P} - \mathbf{P}\mathbf{H}(\mathbf{H}'\mathbf{P}\mathbf{H} + \sigma_v^{s2}\mathbf{I})^{-1}\mathbf{H}'\mathbf{P} \right] \mathbf{F}' + \tilde{\sigma}_\epsilon^2\mathbf{I}$$

and the beliefs evolve according to:

$$\mathcal{E}_t^s \tilde{\mathbf{z}}_{t+1} = \mathbf{F}\mathcal{E}_{t-1}^s \tilde{\mathbf{z}}_t + \mathbf{F}\mathbf{P}\mathbf{H}(\mathbf{H}'\mathbf{P}\mathbf{H} + \sigma_v^{s2}\mathbf{I})^{-1}(x_t - \mathbf{H}'\mathcal{E}_{t-1}^s \tilde{\mathbf{z}}_t) \quad (\text{A.13})$$

$$\mathcal{E}_t^s x_{t+1} = \mathbf{H}'\mathbf{F}\mathcal{E}_{t-1}^s \tilde{\mathbf{z}}_t + \mathbf{H}'\mathbf{F}\mathbf{P}\mathbf{H}(\mathbf{H}'\mathbf{P}\mathbf{H} + \sigma_v^{s2}\mathbf{I})^{-1}(x_t - \mathcal{E}_{t-1}^s x_t) \quad (\text{A.14})$$

Two special cases are of interest. First, when expectations are rational, $\sigma_v^{s2} = 0$, and (A.14) collapses to: $\mathcal{E}_t^s x_{t+1} = \sum_{i=1}^p \lambda_i x_{t+1-i}$, as expected. Second, in the case of an AR(1), the formula simplifies to:

$$\mathcal{E}_t^s x_{t+1} = \lambda(1 - k_t)\mathcal{E}_{t-1}^s x_t + \lambda k_t(x_t - \mathcal{E}_{t-1}^s x_t)$$

where

$$\sigma_{t+1}^2 = (1 - k_t)(\lambda^2 \sigma_t^2 + \tilde{\sigma}_\epsilon^2), \quad 0 \leq k_t = \frac{\lambda^2 \sigma_t^2 + \tilde{\sigma}_\epsilon^2}{\lambda^2 \sigma_t^2 + \sigma_\epsilon^2 + \sigma_v^{s2}}$$

and we obtain lemma 1 as a special case.

A2.3. Equilibrium Exchange Rate

The equilibrium exchange rate is obtained by solving forward the uncovered interest parity condition (A.9): $e_t = \sum_{j=0}^{T-1} \mathcal{E}_t^s (r_{t+j}^* - r_{t+j}) + \mathcal{E}_t^s e_{t+T}$. If we substitute the equation for subjective beliefs (A.14), we obtain:

$$e_t = \bar{e}_t - x_t + \mathbf{H}'(\mathbf{F} - \mathbf{I})^{-1} \mathcal{E}_t^s \tilde{\mathbf{z}}_{t+1} \quad (\text{A.15})$$

where \bar{e}_t , defined as $\lim_{T \rightarrow \infty} \mathcal{E}_t^s e_{t+T}$, is the long run equilibrium exchange rate. Under rational expectations, $\mathcal{E}_t^s \mathbf{x}_{t+1} = \mathcal{E}_t^s \tilde{\mathbf{z}}_{t+1} = \mathbf{F}\tilde{\mathbf{z}}_t = \mathbf{F}\mathbf{x}_t$ where $\mathbf{x}_t = (x_t, x_{t-1}, \dots, x_{t-p+1})$, and the exchange rate follows:

$$e_t^r = \bar{e}_t - x_t + \mathbf{H}'(\mathbf{F} - \mathbf{I})^{-1} \mathbf{F}\mathbf{x}_t \quad (\text{A.16})$$

So that we can rewrite:

$$e_t = e_t^r + \mathbf{H}'(\mathbf{F} - \mathbf{I})^{-1}(\mathcal{E}_t^s \tilde{\mathbf{z}}_{t+1} - \mathbf{F}\mathbf{x}_t)$$

In the case where $p = 1$, the formula simplifies to equation (8):

$$\begin{aligned} e_t &= \bar{e}_t - x_t - \frac{1}{1-\lambda} \mathcal{E}_t^s x_{t+1} \\ &= e_t^r + \frac{1}{1-\lambda} (\mathcal{E}_t^s x_{t+1} - \mathcal{E}_t^s x_{t+1}) \end{aligned}$$

A2.4. Exchange Rate Anomalies

True predictable returns are defined as: $\Lambda_t = \mathcal{E}_t^s e_{t+1} - \mathcal{E}_t e_{t+1}$ where the second expectation is taken with respect to the true process for the forward premium (A.10). Substituting (A.15) and using (A.13) and (A.10), we obtain:

$$\Lambda_t = \left(1 + \mathbf{H}'(\mathbf{F} - \mathbf{I})^{-1} \mathbf{F}\mathbf{P}\mathbf{H}(\mathbf{H}'\mathbf{P}\mathbf{H} + \sigma_v^{s2}\mathbf{I})^{-1} \right) (\mathcal{E}_t^s x_{t+1} - \mathcal{E}_t^s x_{t+1}) \quad (\text{A.17})$$

where $\mathcal{E}_t^s x_{t+1} = \mathbf{H}'\mathbf{F}\mathbf{x}_t$. Predictable excess returns depend linearly upon the misperception in short term interest rates forecasts. In the special case where $p = 1$, this formula simplifies to equation (12):

$$\Lambda_t = \left(1 + \frac{\lambda k}{1-\lambda} \right) (\mathcal{E}_t^s x_{t+1} - \mathcal{E}_t^s x_{t+1})$$

Delayed overshooting: Consider an innovation $\tilde{\epsilon}$ in the forward premium at time t , assuming that $x_{t-j} = 0$ for $j > 0$. At time $t + j$, the forward premium is equal to $\mathbf{F}^j \tilde{\epsilon}$ where $\tilde{\epsilon} = (\tilde{\epsilon}, 0, \dots, 0)'$. Substituting into (A.16), the exchange rate under rational expectations follows

$$e_{t+j}^r = \bar{e}_{t+j} - \mathbf{H}' \left(\mathbf{I} - (\mathbf{F} - \mathbf{I})^{-1} \mathbf{F} \right) \mathbf{F}^j \tilde{\epsilon}$$

which reduces to (13) in the AR(1) case:

$$e_{t+j}^r = \bar{e}_{t+j} - \frac{\lambda^j}{1 - \lambda} \tilde{\epsilon}$$

When beliefs are distorted, the exchange rate can be obtained from (A.15) and (A.13). In the AR(1) case, the formula simplifies. We observe first that

$$\mathcal{E}_{t+j}^s x_{t+j+1} = \lambda (1 - k) \mathcal{E}_{t+j-1}^s x_{t+j} + \lambda^{j+1} k \tilde{\epsilon}$$

Iterating back to $\mathcal{E}_{t-1}^s x_t = 0$, we obtain:

$$\begin{aligned} \mathcal{E}_{t+j}^s x_{t+j+1} &= \sum_{i=0}^j [\lambda (1 - k)]^i \lambda^{j+1-i} k \tilde{\epsilon} \\ &= \lambda^{j+1} \left(1 - (1 - k)^{j+1} \right) \tilde{\epsilon} \end{aligned}$$

The exchange rate then satisfies equation (14):

$$\begin{aligned} e_{t+j} &= \bar{e}_{t+j} - x_t - \frac{1}{1 - \lambda} \mathcal{E}_{t+j}^s x_{t+j+1} \\ &= \bar{e}_{t+j} - \frac{\lambda^j}{1 - \lambda} \left(1 - \lambda (1 - k)^{j+1} \right) \tilde{\epsilon} \end{aligned}$$

The delayed overshooting region D_τ is defined by solving the equation:

$$(1 - k)^\tau = \left(\frac{(1 - \lambda) / \lambda}{1 - \lambda (1 - k)} \right)$$

Forward premium puzzle: We calculate the value of the Fama coefficient as $\beta_0 = \text{cov}(\Delta e_{t+1}, x_t) / \text{var}(x_t)$ for the AR(1) case.³³

$$\begin{aligned} \beta_0 &= \frac{\text{cov}(e_{t+1} - e_t, x_t)}{\text{var}(x_t)} \\ &= 1 - \frac{\text{cov}(\Lambda_t, x_t)}{\text{var}(x_t)} \end{aligned}$$

Using the definition of Λ_t , we see that

$$\begin{aligned} \text{cov}(\Lambda_t, x_t) &= \phi_0 \text{cov}(\mathcal{E}_t x_{t+1} - \mathcal{E}_t^s x_{t+1}, x_t) \\ &= \lambda \phi_0 (1 - k) (\text{var}(x_t) - \lambda \text{cov}(\mathcal{E}_{t-1}^s x_t, x_{t-1})) \end{aligned}$$

where $\phi_0 = (1 - \lambda (1 - k)) / (1 - k)$ and

$$\begin{aligned} \text{cov}(\mathcal{E}_{t-1}^s x_t, x_{t-1}) &= \text{cov}(\lambda (1 - k) \mathcal{E}_{t-2}^s x_{t-1} + \lambda k x_{t-1}, x_{t-1}) \\ &= \lambda^2 (1 - k) \text{cov}(\mathcal{E}_{t-2}^s x_{t-1}, x_{t-1}) + \lambda k \text{var}(x_{t-1}) \end{aligned}$$

so that:

$$\text{cov}(\mathcal{E}_{t-1}^s x_t, x_{t-1}) = (1 - \lambda^2 (1 - k))^{-1} \lambda k \text{var}(x_t)$$

Substituting back, we obtain equation (18):

$$\beta_0 = 1 - \frac{\lambda (1 - \lambda (1 - k)) (1 - k) (1 + \lambda)}{1 - \lambda^2 (1 - k)}$$

Volatility: Assuming that $\bar{e}_t = 0$, the rate of depreciation of the exchange rate satisfies:

$$\begin{aligned} \Delta e_{t+1} &= e_{t+1} - e_t \\ &= x_t + \phi_0 (\mathcal{E}_t^s x_{t+1} - x_{t+1}) \end{aligned}$$

The variance of the spot return is then:

$$\begin{aligned} \text{var}(\Delta e_{t+1}) &= \text{var}(x_t) + \phi_0^2 \text{var}(\mathcal{E}_t^s x_{t+1}) + \phi_0^2 \text{var}(x_{t+1}) \\ &\quad + 2\phi_0 \text{cov}(x_t, \mathcal{E}_t^s x_{t+1}) - 2\phi_0 \text{cov}(x_t, x_{t+1}) - 2\phi_0^2 \text{cov}(\mathcal{E}_t^s x_{t+1}, x_{t+1}) \\ \text{var}(\mathcal{E}_t^s x_{t+1}) &= \text{var}((1 - k) \lambda \mathcal{E}_{t-1}^s x_t + k \lambda x_t) \\ &= \frac{1}{1 - (1 - k)^2 \lambda^2} [k^2 \lambda^2 \text{var}(x_t) + 2(1 - k) k \lambda^3 \text{cov}(\mathcal{E}_{t-1}^s x_t, x_{t-1})] \end{aligned}$$

Similarly,

$$\Lambda_t = \phi_0 [\mathcal{E}_t x_{t+1} - \mathcal{E}_t^s x_{t+1}]$$

and

$$\begin{aligned} \text{var}(\Lambda_t) &= \phi_0^2 (1 - k)^2 \lambda^2 \text{var}(x_t - \mathcal{E}_{t-1}^s x_t) \\ &= \phi_0^2 (1 - k)^2 \lambda^2 [\text{var}(x_t) + \text{var}(\mathcal{E}_{t-1}^s x_t) - 2\lambda \text{cov}(\mathcal{E}_{t-1}^s x_t, x_{t-1})] \end{aligned}$$

Persistence: We calculate the correlation of the depreciation rate as:

$$AC(\Delta e_t) = \text{cov}(\Delta e_{t+1}, \Delta e_t) / \text{var}(\Delta e_t)$$

and

$$\begin{aligned} \text{cov}(\Delta e_{t+1}, \Delta e_t) &= \text{cov}(x_t + \phi_0 (\mathcal{E}_t^s x_{t+1} - x_{t+1}), x_{t-1} + \phi_0 (\mathcal{E}_{t-1}^s x_t - x_t)) \\ &= \lambda \text{var}(x_t) \\ &\quad + \phi_0 \text{cov}(\mathcal{E}_t^s x_{t+1} - x_{t+1}, x_{t-1}) \\ &\quad + \phi_0 \text{cov}(x_t, \mathcal{E}_{t-1}^s x_t - x_t) \\ &\quad + \phi_0^2 \text{cov}(\mathcal{E}_t^s x_{t+1} - x_{t+1}, \mathcal{E}_{t-1}^s x_t - x_t) \end{aligned}$$

Notes

¹See Hodrick (1988) and Lewis (1995) for surveys.

²Empirically, the forward premium puzzle is much more prevalent. The results of Clarida and Gali (1994), Grilli and Roubini (1994) nuance the original results of Eichenbaum and Evans (1995) and indicate that delayed overshooting may not occur for all country pairs.

³See Lewis (1995) and Frankel and Rose (1995) for surveys. Using survey data on exchange rate forecasts, Frankel and Froot (1989) decomposed predictable excess returns into their currency risk premium and expectational error components. Their results indicate that most of the forward premium bias can be attributed to expectational errors.

⁴See Moore and Roche (2002).

⁵Models with learning about a one-time change in regime have been analyzed by Lewis (1989). In these models, investors gradually update their beliefs about the current state of the world, generating systematic forecast errors following a change in regime. These models can explain part of the exchange rate mispredictions, although not the more extreme form where expected depreciation and forward premium move in opposite directions. Models of learning about infrequent regime shifts have very short transitional dynamics. They do not account for the fact that predictable excess returns do not appear to die out over time between infrequent regime switches.

⁶Kaminski (1993) shows that Peso problems can account for part of the forward discount premium.

⁷Tornell (2002) considers such a setup to explain some foreign exchange market anomalies. Hansen, Sargent and Tallarini (1999) use Robust Control to explain macroeconomic anomalies.

⁸See Abel (2002), Barberis, Shleifer and Vishny (1998), Cecchetti, Lam and Mark (2000) and Hong and Stein (1999).

⁹Appendix A2. presents a more complex interest rate process.

¹⁰We do not want to push this interpretation too far. One reason is that the empirical literature on money, output and interest rates tries to separate the exogenous and endogenous components of money shocks. Our univariate representation does not allow for this distinction.

¹¹In the U.S., investors have access to the minutes of the Federal Open Market Committee meetings with a six weeks delay. The full transcript is only available after 3 years.

¹²Recent models include Hong and Stein (1999), Barberis et al. (1998), Cecchetti et al. (2000) and Abel (2002).

¹³This occurs rather fast (see Hamilton (1994)).

¹⁴The last equality follows from the definition of \bar{e}_t .

¹⁵Since the forward premium follows (4), this is the impulse response to the true unconditional shock to the economy.

¹⁶If interest rates follow a white noise process ($\lambda = 0$), there are no forecast misperceptions (since the agents are biased in favor of a white noise to start with) and no predictable currency movements.

¹⁷We thank Michael Moore for encouraging us to look at these two additional puzzles.

¹⁸See Appendix A2..

¹⁹The estimation is performed using Gauss's constrained maximum likelihood (CML) module. In practice, we maximize the likelihood for given values of the initial –unobserved– persistent shock $\mathcal{E}_0 z_1$. We then use the smoother of the Kalman filter to form revised estimates of z_1 , $\mathcal{E}_T z_1$. We iterate back and forth between the Maximum Likelihood and the smoother of the filter until convergence.

²⁰The sample coverage and period is driven by the availability of forecast data for all G-7 countries over that same period. Gourinchas and Tornell (2002) reports results for a longer sample period. The results are unchanged.

²¹For Canada-U.S., the estimated η (0.22) is not significantly different from 0.

²²Because unit root tests lack power, we report Philipps-Perron tests using a longer sample: 1973:01 to 2003:03.

²³In this section, we allow investors to hold subjective beliefs about λ , as well as σ_ϵ^2 , in addition to η . It is easy to show that misperceptions about η are necessary for the puzzle to occur.

²⁴In practice, we estimate θ^s conditional on the true estimate $\mathcal{E}_0 z_1$, estimated in the previous section. In

long enough sample, this is unlikely to matter.

²⁵The Forecasting services that contribute to the Currency Forecaster's Digest are: Predex, Merrill Lynch, Mellon Bank, Harris Trust, Bank of America, Morgan Grenfell, Chase Manhattan, Royal Bank of Canada, Midland Montagu, Generale de Banque, MMS International, Chemical Bank, Union Bank of Switzerland, Multinational Computer Models, Goldman Sachs International, Business International, M. Murenbeeld, and Westpac Bank. The multinational companies that contribute are: General Electric, Du Pont, WR Grace, Allied Signal, Monsanto, Ingersoll-Rand, General-Motors, Data General, Eli Lilly, Aetna, American Express, Johnson and Johnson, Sterling Drug, Firestone, 3M, Union Carbide, Texaco, United Brands, SmithKline Beckman, American National Can, RJ Reynolds, Colgate-Palmolive, Warner-Lambert, Schering-Plough, Quaker Oats, Beatrice Foods, Hercules, Baxter Travenol, and Interpublic Group.

²⁶Ehrbeck and Waldmann (1996) find that models of strategic bias are rejected by the data.

²⁷Results are similar at the 3, 6 and 12 months horizons separately. See Gourinchas and Tornell (2002) for details.

²⁸Formally, $i_t^n = \ln(1 + Y_t^n)$ where Y_t^n is the yield to maturity. Since short term eurodeposits bear no coupons, their duration equals their maturity.

²⁹See Shiller, Campbell and Schoenholtz (1983).

³⁰There are good reasons to view the results of Table 3 with some caution. By analogy with equation (1), we can define the risk premium on the term structure $\zeta_{n,t}^k$ from: $\mathcal{E}_t^s i_{t+n}^k - i_t^n = fp_{n,t}^k - \zeta_{n,t}^k$, where $fp_{n,t}^k \equiv f_{n,t}^k - i_t^n$ is the forward premium on the term structure at horizon n for maturity k . The interpretation is similar to equation (1): the forward premium $fp_{n,t}^k$ provides unbiased estimates of the market forecasts if it is uncorrelated with the risk premium $\zeta_{n,t}^k$. Under the assumption that the consensus forecasts are unbiased estimates of the true market forecasts, we can test whether this restriction holds by running the following OLS regression: $\mathcal{E}_t^s i_{t+n}^k - i_t^n = \alpha + \beta fp_{n,t}^k + \epsilon_{n,t}^k$. Under the null that $\zeta_{n,t}^k$ is uncorrelated with the forward premium, we expect $\beta = 1$. If $\beta < 1$, we would infer that the risk premium and the forward premium are positively correlated. We have run this regression on our sample (results available from the authors) and found systematic evidence that there is a strong risk premium component in the term structure. Froot (1989) obtains similar results using different market forecasts.

³¹Empirical estimates for the delayed overshooting horizon (τ) are from Eichenbaum and Evans (1995),

Table Ib, line 9. Estimates for β are from Gourinchas and Tornell (2002), table 1. An estimate of $var(\Lambda)$ is constructed from the variance of the fitted values of a regression of realized excess returns $e_{t+3} - e_t - x_t$ on x_t . The coefficient from this regression is estimated as the sum of cross-correlations between Δe_{t+3} , Δe_{t+2} , Δe_{t+1} and x_t , times the ratio of the standard deviation of Δe_{t+1} and the standard deviation of x_t . The sample period is 1974:1 to 1999:12.

³²w.l.o.g. we can assume that all the innovations have the same variance since each innovation is scaled.

³³The general case can be solved similarly. However, there is no closed-form solution.

Country pair	$AR(1)$			$AR(2)$			Z_t
	λ	η	Log-lik.	$\Sigma\lambda$	η	Log-lik.	
U.S.-Canada	0.95 (0.03)	0.22 (0.23)	-0.28	0.95 (0.03)	0	-0.22	-3.33 [0.01]
U.S.-France	0.96 (0.04)	0	-0.44	0.96 (0.04)	0	-0.43	-4.94 [0.00]
U.S.-Germany	0.99 (0.01)	0	0.10	0.99 (0.01)	0	0.14	-2.51 [0.11]
U.S.-Italy	0.98 (0.02)	0	-0.48	0.98 (0.02)	0	-0.47	-2.16 [0.22]
U.S.-Japan	1.01 (0.01)	0	0.19	1.00 (0.01)	0	0.21	-2.92 [0.04]
U.S.-U.K.	0.98 (0.02)	0	-0.27	0.98 (0.01)	0	-0.26	-3.99 [0.00]
	$AR(3)$			$AR(4)$			
	$\Sigma\lambda$	η	Log-lik.	$\Sigma\lambda$	η	Log-lik.	
U.S.-Canada	0.92 (0.04)	0	-0.22	0.91 (0.04)	0	-0.22	
U.S.-France	0.96 (0.04)	0	-0.43	0.98 (0.03)	0	-0.41	
U.S.-Germany	0.99 (0.01)	0	0.13	0.99 (0.01)	0	0.17	
U.S.-Italy	0.99 (0.01)	0	-0.47	0.99 (0.02)	0	-0.47	
U.S.-Japan	1.00 (0.01)	0	0.22	1.00 (0.01)	0	0.22	
U.S.-U.K.	0.97 (0.02)	0	-0.25	0.99 (0.01)	0	-0.24	

Table 1: Maximum Likelihood Estimates of the State-Space Representation of the 3-months Eurorates. Z_t reports the Phillips-Perron tests of a unit root in the forward premium. p-values in brackets.

Country pair	$AR(1)$			$AR(2)$			Obs
	λ^s	η^s	Log-lik.	$\Sigma\lambda^s$	η^s	Log-lik.	
U.S.-Canada	0.97 (0.004)	2.81 (0.59)	0.312	0.95 (0.007)	0.67 (0.07)	0.331	312
U.S.-France	0.97 (0.003)	0.94 (0.19)	-0.286	0.99 (0.001)	7.97 (0.84)	-0.222	237
U.S.-Germany	0.98 (0.002)	0.08 (0.07)	0.273	0.99 (0.001)	7.21 (1.51)	0.434	312
U.S.-Italy	0.99 (0.004)	1.36 (0.27)	-0.415	1.00 (0.001)	9.68 (3.03)	-0.368	216
U.S.-Japan	0.98 (0.002)	0.20 (0.07)	0.530	0.99 (0.001)	3.11 (1.08)	0.600	306
U.S.-U.K.	0.98 (0.002)	0.57 (0.06)	0.193	0.99 (0.001)	8.58 (2.34)	0.279	306
	$AR(3)$			$AR(4)$			
	$\Sigma\lambda^s$	η^s	Log-lik.	$\Sigma\lambda^s$	η^s	Log-lik.	
U.S.-Canada	0.92 (0.007)	0.11 (0.04)	0.348	0.90 (0.008)	0.03 0.03	0.373	
U.S.-France	0.99 (0.001)	17.94 (6.37)	0.011	0.98 (0.001)	4.08 (0.67)	0.032	
U.S.-Germany	0.99 (0.002)	4.55 (1.49)	0.533	0.99 (0.002)	2.82 (1.13)	0.545	
U.S.-Italy	0.99 (0.002)	2.67 (1.24)	-0.342	0.99 (0.001)	0.09 (0.08)	-0.216	
U.S.-Japan	0.98 (0.001)	1.49 (0.38)	0.676	0.98 (0.003)	0.54 (0.22)	0.686	
U.S.-U.K.	0.99 (0.001)	29.51 (5.55)	0.371	0.97 (0.003)	0.85 (0.36)	0.339	

Table 2: Maximum Likelihood Estimates of Subjective Beliefs using Consensus Forecasts

	$AR(1)$			$AR(2)$			Obs
	λ^s	η^s	Log-lik.	$\Sigma\lambda^s$	η^s	Log-lik.	
U.S.-Canada	0.95 (0.006)	0.36 (0.37)	0.970	0.95 (0.006)	0.26 (0.15)	0.970	220
U.S.-France	0.93 (0.01)	0	0.225	0.95 (0.002)	0.59 (0.13)	0.225	150
U.S.-Germany	0.98 (0.003)	0	0.303	0.99 (0.001)	2.05 (0.91)	0.742	220
U.S.-Italy	0.97 (0.003)	0	0.221	0.98 (0.002)	1.06 (0.31)	0.285	220
U.S.-Japan	0.99 (0.004)	0	1.535	0.99 (0.001)	1.06 (0.23)	1.692	150
U.S.-U.K.	0.97 (0.002)	0	0.475	0.98 (0.002)	0.91 (0.44)	0.655	220
	$AR(3)$			$AR(4)$			
	$\Sigma\lambda^s$	η^s	Log-lik.	$\Sigma\lambda^s$	η^s	Log-lik.	
U.S.-Canada	0.96 (0.004)	0.18 (0.06)	1.052	0.95 (0.004)	0	1.055	
U.S.-France	0.96 (0.004)	2.63 (0.77)	0.293	0.97 (0.004)	4.44 (2.15)	0.318	
U.S.-Germany	0.99 (0.001)	1.24 (0.56)	0.808	0.99 (0.001)	0.94 (0.30)	0.837	
U.S.-Italy	0.97 (0.003)	0.04 (0.11)	0.252	0.98 (0.002)	1.31 (0.56)	0.305	
U.S.-Japan	0.99 (0.001)	1.18 (0.51)	1.710	0.99 (0.001)	2.19 (0.32)	1.830	
U.S.-U.K.	0.98 (0.002)	0.72 (0.34)	0.655	0.98 (0.002)	0.68 (0.26)	0.671	

Table 3: Maximum Likelihood Estimation of the Subjective Beliefs using Forward Rates

		k	τ	β	$\frac{var(\Delta e)}{var(x)}$	$\frac{var(\Delta)}{var(x)}$	$AC(\Delta e_t)$
U.S.-Canada	data		NA	-0.87	30.02	4.69	-0.06
	implied	0.68 (0.02)	2.25 (0.16)	0.42 (0.03)	20.15 (2.95)	1.89 (0.34)	0.25 (0.02)
U.S.-France	data		37.47	-0.16	53.13	1.57	0.01
	implied	0.29 (0.01)	9.41 (0.38)	-0.36 (0.02)	28.10 (2.36)	14.67 (0.79)	0.68 (0.01)
U.S.-Germany	data		35.30	-0.50	71.08	2.56	0.04
	implied	0.31 (0.03)	10.22 (0.97)	-0.35 (0.05)	47.50 (4.27)	23.39 (1.49)	0.67 (0.03)
U.S.-Italy	data		37.99	1.41	76.98	0.49	0.08
	implied	0.27 (0.04)	13.45 (1.91)	-0.43 (0.07)	80.40 (28.90)	43.31 (13.76)	0.72 (0.04)
U.S.-Japan	data		24.65	-2.98	109.65	15.74	0.08
	implied	0.42 (0.05)	6.92 (1.09)	-0.12 (0.10)	57.44 (7.82)	19.05 (2.51)	0.55 (0.05)
U.S.-U.K.	data		38.87	-1.25	76.52	5.36	0.08
	implied	0.28 (0.03)	10.3 (1.26)	-0.38 (0.06)	34.49 (4.31)	18.31 (1.57)	0.69 (0.03)

Table 4: Accounting for the Empirical Puzzles

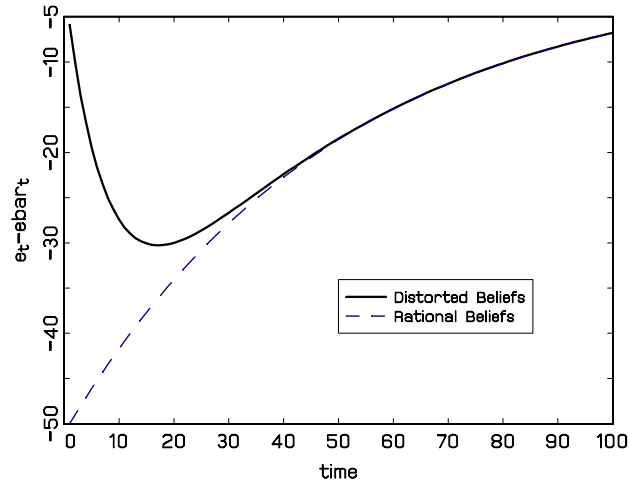


Figure 1: Delayed Overshooting Response to an Unanticipated Increase in Domestic Interest Rate. $k = 0.1$ and $\lambda = 0.98$.

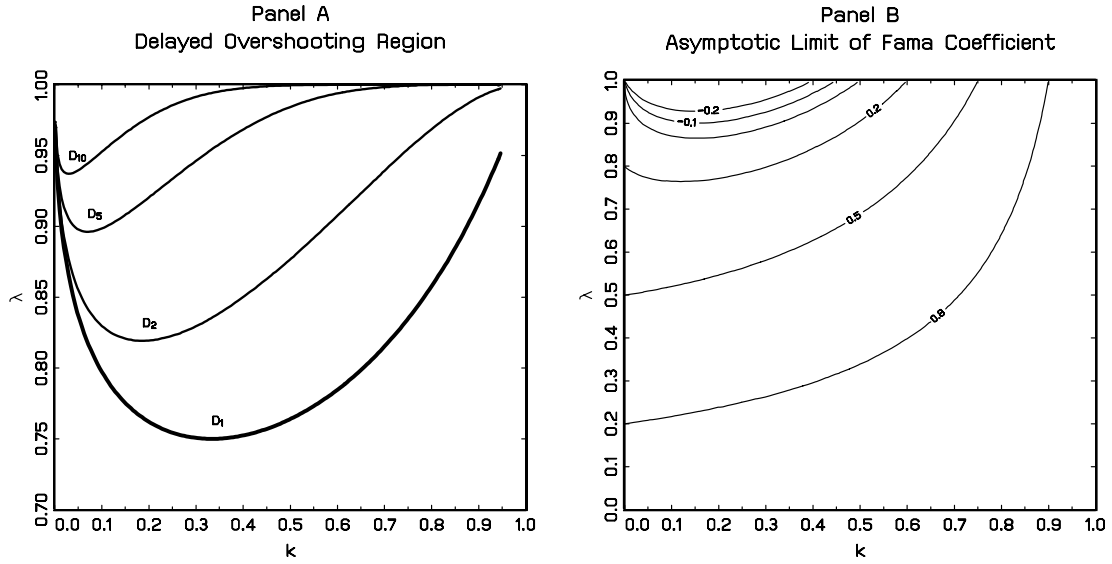


Figure 2: Delayed Overshooting and Forward Premium Puzzle Regions.

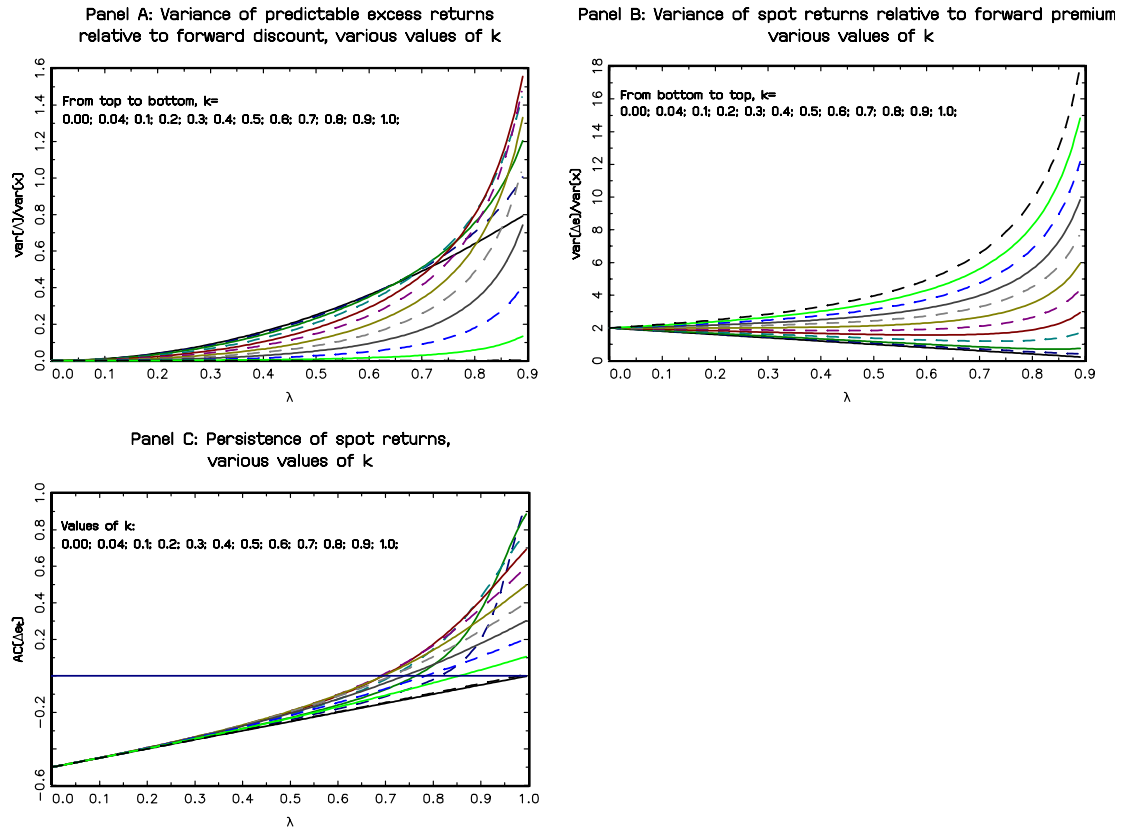


Figure 3: Volatility and Persistence Puzzles.

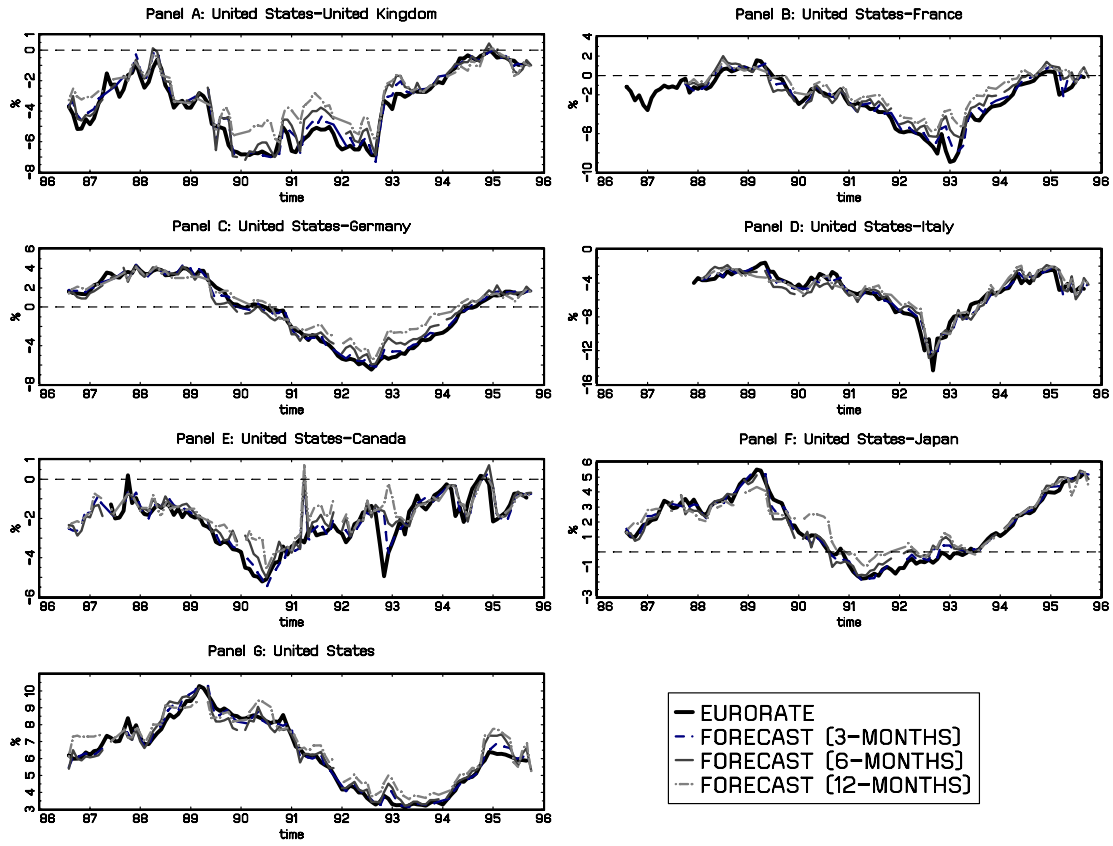


Figure 4: 3-months Eurorates and 3, 6 and 12 months forecasts, against the U.S. dollar (panel A-F) and 3-months eurodollar (panel G).